

**A UNIFIED TRANSLITERATION SYSTEM FOR MAINLAND SOUTHEAST ASIAN SCRIPTS:
GUIDE AND TABLES**

Trent Walker

The following is an online appendix to “A Unified Transliteration System for Mainland Southeast Asian Scripts: Principles, Inventory, and Examples” (<https://doi.org/10.69486/114.1.2026.6a>). It provides the full set of 25 tables necessary for implementing the transliteration system as described in the article. To make the tables easier to navigate, this appendix begins with a short guide that introduces the structure of the tables, the logic behind this structure, and salient points for applying them in the transliteration of Southeast Asian texts.

Guide to the Tables

The tables include support for 15 different scripts, abbreviated as follows:

1. Kh. Chr. = Khmer script, Chrieng (“slanted”) style
2. Kh. M. = Khmer script, Mul (“rounded”) style
3. K. Thai = Khom Thai script, Banchong (“upright”) style
4. Thai = modern Thai script
5. Lao = modern Lao script
6. Lao B. = Lao Buhan/Thai Noi script
7. T. Lao = Tham script, Lao/Isan variant
8. Lanna = Tham script, Lanna variant
9. Khuen = Tham script, Khuen variant
10. T. Lue = Tham script, Lue variant
11. Lue = new Tai Lue (Dai Lue) script
12. Lik = Lik Tai script, with the Tai Ahom variant used as a reference (provisional)
13. Shan = traditional and modern Shan script
14. Mon = Mon script, with Burmese Mon, rather than Thai Mon, used as a reference
15. Brm. = Burmese/Myanmar script

To make the tables easier to navigate, the scripts are arranged to show a spectrum from the twin poles of Khmer/Khom on the left and Mon/Burmese scripts on right. This leaves most of the Tai scripts in the middle. Those closer to Khmer in both current form and paleographic development, namely Thai and Lao, appear toward the left, while those closer to Mon/Burmese, namely Lik Tai and Shan, are placed on the right. The Tham family of scripts, which draws influences from both Khmer/Khom and Mon/Burmese, occupies the middle of the table; Lao Buhan and the Lik Tai family of scripts, which occasionally borrow letterforms from the Tham family, flank its left and right sides, respectively. For all 15 scripts, modern fonts are used for the sake of convenience. This means that not every symbol found in manuscripts and inscriptions can be adequately represented with a Unicode symbol in the selected fonts. For these exceptional cases, I have either combined multiple fonts (sometimes from multiple different scripts) or inserted a custom-designed image to provide a visual match for the grapheme in question. That said, the tables are not intended as comprehensive guides to Southeast Asian paleography but rather as streamlined tools for transliteration.

TABLE 1: Consonants

The first table lists 78 consonants that may appear as initials, medials, or finals. They are presented in dictionary order (i.e., customary Indic order: *ka, kha, ga, gha, na*, etc.), following the norm for scripts across the region. The consonants include the standard Indic ones, modified versions of these to fit Southeast Asian phonologies, and a handful of diacritical consonants, which appear only in medial or final position. A few special glyphs, namely for *pā* and *nā* in selected scripts, have been included for the sake of comparison; the diverse forms for *ññ* and *ss* represent separate orthographic variants. For the standard Indic consonants, many scripts include a complete or partial set of subscript consonants, which appear in smaller script on the right side of each box.

TABLE 2: Independent Vowels

The second table provides letterforms for 21 different independent vowels—vowels that most commonly appear in an initial position in a word or form a word by themselves, not tied to any consonant. Special rules for Khmer, Khom, Tham, Mon, and Burmese scripts concerning

independent vowels and dependent vowels attached to the letter *a* are discussed at length in the main article (“Principles, Inventory, and Examples”). The remaining scripts, including Lao, Lue, and Shan, do not have specific letter forms for independent vowels. Forms for *a* and *ā* in these scripts are included in **TABLE 2** for the sake of comparison. When the glottal-stop/vowel-support consonant *a* in initial position is combined with dependent vowels in these scripts, they are simply transliterated as *i*, *ī*, *u*, etc. (e.g., Thai *o* ๑๐๓, “to take”; Shan *ik* ၵိၵ်ႈ, a conjunction). Among the remaining scripts discussed here, only Thai has a full set of vocalic consonants (*r*, *ṛ*, *l*, *ḷ*), traditionally treated as independent vowels. The vocalic consonants do not generate any additional ambiguities in transliteration, as the present system follows the ISO standard in using the ring below symbol, whether in initial or medial position.

TABLES 3.1–3.17: Dependent Vowels

These 17 tables provide letterforms for 209 different dependent vowels and vowel diacritics, both alone and in combination with semivowel finals. Each is appended to the dummy syllable *ka*. The semivowel finals *-y* and *-v*—along with various diacritics, such as the *anusvāra* [ṁ] and the *visarga* [ḥ]—are included as they match traditional Southeast Asian definitions of dependent vowels. The number of vowels is also increased by the inclusion of orthographically meaningful variants for *æ* and *æa*. Allographs—variants that only affect the shape of characters as opposed to the spelling of words—are excluded. These include glyphs of varying heights for *ā*, *ī*, *ṛ*, etc. when paired with specific consonants; see Hundius (1990: 179–189, 222–244) for abundant examples of such allographs in the Lanna variety of Tham script.

As noted above, the 209 dependent vowels are divided into 17 different groups: *a*, *ṛ*, *ā*, *i*, *ī*, *u*, *ṛ*, *e*, *æ*, *ai*, *o*, *au*, *æ*, *æa*, *ia*, *ua*, and *ui*. Each group (except *ui*) coalesces around a particular phonological system: *a* → /a/, *ṛ* → /a/, *ā* → /a:/, *i* → /i/, *ī* → /i:/, *u* → /u/, *ṛ* → /ru/, *e* → /e/, *æ* → /ɛ/, *ai* → /aj/, *o* → /o/, *au* → /aw/, *æ* → /ɤ/, *æa* → /ua/, *ia* → /ia/, and *ua* → /ua/. The vowels transliterated with a single letter (*a*, *ṛ*, *ā*, *i*, *ī*, *u*, *ṛ*, *e*, *æ*, *o*, *æ*) are usually monophthongs, though diphthongized pronunciations of many of these have developed for Khmer and Mon. The vowels transliterated with two letters (*ai*, *au*, *æa*, *ia*, and *ua*; again, *ui* is an exception) are more often pronounced as diphthongs, though in several languages some of these have reverted to monophthongs. The forms for *ia* and *ua* are originally based on the subscripts for the semivowels

ya and va, respectively; when these consonantal sounds are operative, subscript ya and va should be transliterated as such rather than by the simplest forms of ia or ua. By and large, written vowel forms belong exclusively to one group, though a few overlapping forms exist for æ and æa. These overlapping forms do not pose any obstacles for the one-way reversibility of the system, though they require discernment on the part of the reader to determine which phonological principles should be applied.

The Roman-script graphemes selected for transliteration are intended to be modular and reflective of the original scripts. Hence the series *i*, *ī*, *î*, *ï* mirrors the way the basic *i* sign is modified into long *ī* with a single line, into *î* with a single dot, and into *ï* with a combination of added lines or dots. The same logic applies to the series *æ*, *ǣ* (long *œ*), *ǿ* (*œ*+*ï*), and *ǫ̇* (*œ*+*ï*). The basic vowel *œ* is represented by a digraph that includes the symbol for *i* on top. If the *i* portion of the digraph is shifted to *ī*, then the transliteration shifts to *ǣ*. If the *i* portion is shifted to *î*, then the transliteration becomes *ǿ*; the use of the *ï* character results in *ǫ̇*. If the *i* portion is dropped altogether, then the *œ* is replaced by *ø* (e.g., Thai *kløq* เกล็ด, “friend”; cf. Thai *kæt* เกิด, “to be born”). When various diacritics and semivowels are added to these basic combinations, a similar method applies. For each of the written vowels composed of multiple underlying vowel characters, diacritics, or semivowels, I provide a simple equation in the transliteration to help users appreciate the rationale behind the formation of each of the forms. This is designed to make the system easier to remember without recourse to the tables.

TABLE 4: Tone Marks

This table features tone marks placed on or after the dummy syllable *kā*. The first row of the table shows *kā* without any added tone marker. Rows 2–7 indicate tones whose transliterations involve diacritics placed directly above first vowel letter of the modified syllable (e.g., Lao *kā́* ກ້, “to dare”). The remaining rows, used only for Burmese script and late 19th-century through mid-20th-century Shan script, provide transliterations for tones that appear at the end of the syllable instead (e.g., Shan *cam*. cṵ, a nominative/accusative marker; cf. modern Shan *cām*. coṵ).

Apart from Burmese and older forms of Shan, in which tone marks appear at the end of the syllable they modify, tone marks in transliteration are otherwise placed directly on top of the vowels of the syllables they modify. When the vowel in question is written as a diphthong, the

tone mark always appears on the first of the two vowels in transliteration (e.g., Thai *hái* ให้, “to give”; Tham *kìav* 𑜋𑜊𑜍, “muscle cramp”).

TABLE 5: Other Diacritics

A handful of diacritics are not covered by the previous tables. These symbols—not easily grouped with consonants, vowels, or tone marks—take on a variety of functions. Two of the most prominent symbols in this system—the macron below and the equal sign for marking special uses of subscripts—are discussed at length in the main article, along with the interpunct for the *virāma*; the bullet for the *daṇḍaghāt*; the apostrophe for the Sanskrit *avagraha*; the macron above for the Thai *yāmākkār*; the tilde, the double tilde, and the ampersand for different repeating marks; and the hyphen for bridging various forms of hiatus. Among those not discussed in “Principles, Inventory, and Examples” are two symbols that are reused from the tone marks table for non-tonal purposes. The grave accent stands for the Khmer *rassasaññā*, a vowel shortening mark identical in appearance to the first tone mark in Thai, Lao, and Tham scripts. This diacritic is typically placed over the final consonant (e.g., Modern Khmer *kāt* កាត់, “to cut”), but its position may be adjusted depending on the document in question (e.g., Middle Khmer *kāt* កាត់; similar spellings also appear with the *saṃyog saññā* instead, e.g., *kāt* កាត់). The Khmer syllable modifier known as the *kākapād*, identical in form to the fourth tone marker in Thai, is likewise transliterated with the hook above, placed above the consonant in question (e.g., Modern Khmer *cā* ចាំ, a polite response particle; Middle Khmer *tām* *tāe* តាមតែ, “in exact accordance”).

TABLE 6: Numerals

Southeast Asian writing systems often make use of multiple sets of numerals. This table matches the Hindu–Arabic digits 0–9 with their equivalents in 18 different systems. *Lekh attā* are a set of numerals formerly used for mathematics and astrology in Cambodia and Thailand. In the Tham-script sphere, there are two main sets of numerals, a first set used primarily in Buddhist texts (*lekh nai dhamm*) and a second set, closely related to Burmese numerals, that are used in other applications. When a manuscript uses multiple sets of numerals for a given script—particularly

common in Tham-script contexts—underlining can be used to distinguish between each set, as suggested in the table headings.

TABLE 7: Abbreviations

Several scripts, notably Tham and Burmese, use a variety of special characters to abbreviate common words; these abbreviations, including several combined forms, are listed in this table. Tham script also includes a variety of special compressed spellings (Hundius 1990: 244–245). Since most of these can be accurately transliterated with the equal sign (e.g., Tham *pa=mī* နှိ, “does not exist”), or with the equal sign plus the degree symbol for abbreviated forms (e.g., Tham *ḍī=la°* နှိ, i.e., *ḍī hlī* နှိ, “very good”), these orthographies are not included in the table to make room for abbreviations not otherwise accounted for in the present system. These include distinct letterforms commonly used as abbreviations in Burmese and Tham scripts as well as a few of the possible logographic forms in Tham. As more such abbreviated forms are identified by researchers, this portion of the system is ready for expansion to include additional forms preceded by the backtick (`) as a generic special character mark.

TABLE 8: Punctuation

This table combines several different kinds of features found in Southeast Asian documents. These include various forms of traditional punctuation marks, including variations of the Indic *daṇḍa*, elision markers known as *peyyāla*, and specialized markers for beginnings and endings. Many of these marks can be combined in complex ways; the table does not list such combinations exhaustively but presents some of the most common types. Punctuation borrowed from Roman conventions, including periods, commas, and question marks as used in modern Lao, are not listed in the table; such Roman-script marks should be included verbatim in the transliteration. The table also includes specific features commonly found in manuscripts and inscriptions, including spaces in the original (_), line breaks (/), space fillers (#), bracketed lists, section marks, and other symbols relevant for layout and formatting.

TABLE 9: Editorial Marks

The final table provides a list of symbols used in the editorial markup of a transliterated text. Many of these concern the complexities of word shape, as highlighted above, but others are dedicated for broader editorial needs such as correction, addition, and deletion. A few symbols and combinations thereof are specifically set aside for recording traditional editorial interventions made by scribes, proofreaders, and annotators, including those found in interlinear and marginal notes. This table only provides a basic guideline for some of the key editorial interventions that transliterators of Southeast Asian texts will need to make; some documents may require a more complex editorial apparatus. The foundational set of editorial marks listed here are selected with two key principles in mind: (1) consonance with existing scholarly norms for diplomatic transcriptions and critical editions, and (2) non-overlap with the panoply of other symbols required for the transliteration of Southeast Asian texts.

ONLINE APPENDIX

TABLE 1 — Consonants

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
k	ក្រ	ក	ក	ก	ກ	ກ	ກ	ກ	ກ	ກ	ກ	ກ	ກ	က	က
k̥ Diacritical final k							◌◌								
kh	ខ	ខ	ខ	ข	ຂ	ຂ	ခ	ခ	ခ	ခ	ခ	ခ	ခ	ခ	ခ
kh				ข				ခ	ခ	ခ			ခ		
g	ဂ	ဂ	ဂ	ค	ຄ	ຄ	ဂ	ဂ	ဂ	ဂ	ဂ		ဂ	ဂ	ဂ
ḡ				ค				ဂ	ဂ	ဂ	ဂ				
gh	ဃ	ဃ	ဃ	ฆ	ພ		ဃ	ဃ	ဃ	ဃ			ဃ	ယ	ယ
ṇ	ဏ	ဏ	ဏ	ง	ງ	ງ	ဏ	ဏ	ဏ	ဏ	ဏ	ဏ	ဏ	ဏ	ဏ
ṇ̃ Diacritical medial ṇ	န	န					န	န	န	န				န	န
ṇ̣ Small diac. med. ṇ								န	န	န					
ṇ̇	န	န													

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
c	င	င	င	จ	ຈ	ຈ	င	င	င	င	င	၈	င		င
ch	စ	စ	စ	ฉ	ຂ		ဆ	ဆ	ဆ	ဆ			ဆ	ဆ	ဆ
j	ည	ည	ည	ช	ຊ	ຊ	ည	ဂ	ဂ	ဂ	ဂ		ဂ	ဂ	ဂ
j̣				ช		ຊ	ဂ	ဂ	ဂ	ဂ	ဂ		ဂ		
jh	လ	လ	လ	ฉ	ຊ		ဆ	ဆ	ဆ	ဆ			ဆ	ဆ	ဆ
ñ	ဣ	ဣ	ဣ	ญ	ຍ		ယ	ယ	ယ	ယ		ဖ	ယ	ယ	ယ
ṇ̃	ဣ / ဣ	ဣ / ဣ													
ññ Special glyph	ဣ	ဣ	ဣ				ယ	ယ	ယ	ယ				ယ	ယ
ñṇ̃ Mon ññ														ယ	
ñṇ̃ Alt. Mon ññ														ယ	
ṭ	တ	တ	တ	ต	ຕ		တ	တ	တ	တ			တ	တ	တ
ṭ̣				ต											

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
ṭh	ṭ ^ṭ	ṭ ^ṭ	ṭ ^ṭ	ṭ	ṭ	ṭ	ṭ	ṭ	ṭ	ṭ			ṭ	ṭ	ṭ
ḍ	ḍ ^ḍ	ḍ ^ḍ	ḍ ^ḍ	ḍ	ḍ	ḍ	ḍ	ḍ	ḍ	ḍ	ḍ		ḍ	ḍ	ḍ
ḍ̣											ḍ̣				
ḍh	ḍ ^ḍ	ḍ ^ḍ	ḍ ^ḍ	ḍ	ḍ		ḍ	ḍ	ḍ	ḍ			ḍ	ḍ	ḍ
ṇ	ṇ ^ṇ	ṇ ^ṇ	ṇ ^ṇ	ṇ	ṇ		ṇ	ṇ	ṇ	ṇ			ṇ	ṇ	ṇ
t	t ^t	t ^t	t ^t	t	t	t	t	t	t	t	t	t	t	t	t
ṭ				ṭ	ṭ	ṭ									
th	th th	th th	th th	th	th	th	th	th	th	th	th	th	th	th	th
d	d ^d	d ^d	d ^d	d	d	d	d	d	d	d	d		d	d	d
dh	dh ^{dh}	dh ^{dh}	dh ^{dh}	dh	dh		dh	dh	dh	dh	dh		dh	dh	dh
n	n ⁿ	n ⁿ	n ⁿ	n	n	n	n	n	n	n	n	n	n	n	n
nā							nā	nā	nā	nā					
Special glyph															
ñ	ñ ^ñ	ñ ^ñ													

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
p	ပု	ပု	ပု	ပ	ပ	ပ	ပ	ပ	ပ	ပ	ပ	ပ	ပ	ပ	ပ
pā	ပာ	ပာ	ပာ												
Special glyph															
p̂	ပီ	ပီ									ပီ				
p̣			ပ	ပ	ပ	ပ	ပ	ပ	ပ	ပ	ပ				
p̥	ပီ / ပု	ပီ / ပု													
ph	ဖ	ဖ	ဖ	ဖ	ဖ	ဖ	ဖ	ဖ	ဖ	ဖ	ဖ	ဖ	ဖ	ဖ	ဖ
p̄h			ဖ	ဖ	ဖ	ဖ	ဖ	ဖ	ဖ	ဖ	ဖ		ဖ		
b	ပ	ပ	ပ	ပ	ပ	ပ	ပ	ပ	ပ	ပ	ပ		ပ	ပ	ပ
ḅ			ပ	ပ	ပ	ပ	ပ	ပ	ပ	ပ	ပ			ပ	
b̥														ပ	
bh	ပ	ပ	ပ	ပ	ပ		ပ	ပ	ပ	ပ	ပ		ပ	ပ	ပ
m	မ	မ	မ	မ	မ	မ	မ	မ	မ	မ	မ	မ	မ	မ	မ
ṃ	မီ / မု	မီ / မု													

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
y	យ ្យ	យ ្យ	ຢ ຢ	ย	ຍ	ຍ	ຍ ្យ	ຍ ្យ	ຍ ្យ	ຍ ្យ	ຍ		ယ	ယ ្យ	ယ ្យ
ÿ												ᵛ ᵛ	ᵛ ᵛ	ᵛ ᵛ	
Diacritical final y															
ŷ	យ̣ / យ̣	យ̣ / យ̣													
ȳ					ຢ	ຍ	ຍ ្យ	ຍ ្យ	ຍ ្យ	ຍ ្យ	ຍ				
Inital/medial Tai *y															
Ț							ຍ ្យ	ຍ ្យ	ຍ ្យ	ຍ ្យ					
Final subscript y															
r	រ ្រ	រ ្រ	ຣ ຣ	ร	ຣ	ຣ	ຣ ្រ	ຣ ្រ	ຣ ្រ	ຣ ្រ	ຣ		ຣ	ຣ ្រ	ຣ ្រ
ṛ								ᵛ							
Subscript final r															
ř	ຣ̣	ຣ̣	ຣ̣			ຣ̣	ຣ̣	ຣ̣	ຣ̣	ຣ̣		ᵛ		ᵛ	ᵛ
Repha/diacritical r															
rr			ຣຣຣ	ຣຣຣ											
Ra ḥṇ, Thai ř															
rř	ຣຣ̣	ຣຣ̣	ຣຣ̣												
ṛ̌	ຣ̣ / ຣ̣	ຣ̣ / ຣ̣													
l	လ ်	လ ်	လ ်	ล	ล	ล ်	ล ်	ล ်	ล ်	ล ်	ล	ล	ล	ล ်	ล ်

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
V	វ៌	វ៌ / វ៌	ว	ว	ວ	ວ	ວ	ວ	ວ	ວ	ວ	ວ	ວ	ວ	ວ
Ṽ Diacritical final v								◌ [ꠘ]	◌ [ꠘ]	◌ [ꠘ]				◌	
Ṽ	វ៌ / វ៌	វ៌ / វ៌													
Ś	ศ	ศ	ศ	ศ			ศ	ศ	ศ	ศ				ศ	ศ
Ṣ	ษ	ษ	ษ	ษ			ษ	ษ	ษ	ษ				ษ	ษ
Ṣ̣ Lao B. subscript ṣ						ṣ									
S	ສ	ສ	ส	ส	ສ	ສ	ສ	ສ	ສ	ສ	ສ	ສ	ສ	ສ	ສ
SS Special glyph							ສ	ສ	ສ	ສ				ສ	ສ
SṢ Thai Mon ss														ສ	
Ṣ̣	ສ / ស	ສ / ស													
h	ហ	ហ	ห	ห	ຫ	ຫ	ຫ	ຫ	ຫ	ຫ	ຫ	ຫ	ຫ	ຫ	ຫ
ĥ	ហ / ហ	ហ / ហ													
l	ล	ล	ล	ล			ล	ล	ล	ล			ล	ล	ล

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
ᵀ	ᵀ	ᵀ													
a / ^a (see article)	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ
^a â (^ over vowel)	ᵀ / ᵀ	ᵀ / ᵀ									ᵀ				
ᵀ				ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ					

TABLE 2 — Independent Vowels

Full sets in Khmer, Khom, Tham, Mon, and Burmese only; not applicable in Thai, Lao, Lue, Lik Tai, and Shan

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
a	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ
ā (long ā)	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ
i	ᵀ	ᵀ	ᵀ				ᵀ	ᵀ	ᵀ	ᵀ				ᵀ	ᵀ
ī (long ī)	ᵀ	ᵀ	ᵀ				ᵀ	ᵀ	ᵀ	ᵀ				ᵀ	ᵀ
ᵀ (modified i)			ᵀ												
ī (long i)			ᵀ												
u	ᵀ	ᵀ	ᵀ				ᵀ	ᵀ	ᵀ	ᵀ				ᵀ	ᵀ

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
ū (long ū)	ຸ	ຸ	ຸ				ຸ	ຸ	ຸ	ຸ				ຸ	ຸ
ū̄ (modified ū/ūv)	ຸ	ຸ													
û (u+â/uk)	ຸ	ຸ													
r̄	ຸ	ຸ	ຸ	ຸ											ຸ
r̄̄ (long r̄)	ຸ	ຸ	ຸ	ຸ			ຸ	ຸ	ຸ	ຸ					ຸ
l̄	ຸ	ຸ	ຸ	ຸ											ຸ
l̄̄ (long l̄)	ຸ	ຸ	ຸ	ຸ				ຸ	ຸ	ຸ					ຸ
e	ຸ	ຸ	ຸ				ຸ	ຸ	ຸ	ຸ				ຸ	ຸ
ẹ (modified e)			ຸ												
ai	ຸ	ຸ	ຸ												ຸ
o	ຸ	ຸ	ຸ				ຸ	ຸ	ຸ	ຸ				ຸ	ຸ
ō (modified o/u)	ຸ	ຸ													
au	ຸ	ຸ	ຸ				ຸ	ຸ	ຸ	ຸ					ຸ
âu (au+â)							ຸ	ຸ	ຸ	ຸ					

TABLE 3 — Dependent Vowels

TABLE 3.1 — A

Usually pronounced /a/; other pronunciations include /o/, /ěə/, /ǝə/, and /ɛ/ (for *ay* and *av* diphthongs, see tables for *ai* and *au*)

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
a (no written vowel)	𑀓	𑀓	𑀓	𑀓	𑀓	𑀓	𑀓	𑀓	𑀓	𑀓	𑀓	𑀓	𑀓	𑀓	𑀓
ä (a+̣) (yugalabindu; inherent a vowel)	𑀓̣	𑀓̣													
â (a+̂) (addhacand; inherent a/[o] vowel)	𑀓̂	𑀓̂			𑀓̂	𑀓̂	𑀓̂	𑀓̂	𑀓̂	𑀓̂					
ǣ (a+̃) (musikadāni; inherent vowel)			𑀓̃	𑀓̃											
aṁ (a+ṁ) (anusvara/niggahīta)	𑀓̣̣	𑀓̣̣	𑀓̣̣	𑀓̣̣	𑀓̣̣	𑀓̣̣	𑀓̣̣	𑀓̣̣	𑀓̣̣	𑀓̣̣		𑀓̣̣	𑀓̣̣	𑀓̣̣	𑀓̣̣
aṃ̇ (aṁ+̣) (candrabindu, Sanskrit only)	𑀓̣̣̣	𑀓̣̣̣													
aḥ (a+ḥ) (visarga)	𑀓̣̣̣̣	𑀓̣̣̣̣	𑀓̣̣̣̣	𑀓̣̣̣̣	𑀓̣̣̣̣	𑀓̣̣̣̣	𑀓̣̣̣̣	𑀓̣̣̣̣	𑀓̣̣̣̣	𑀓̣̣̣̣	𑀓̣̣̣̣	𑀓̣̣̣̣		𑀓̣̣̣̣	𑀓̣̣̣̣
ă (a+̣̣) (saṃyog saññā/mái hñ ākās/ mái jāt; short a or final ăk/-k)	𑀓̣̣̣̣̣	𑀓̣̣̣̣̣	𑀓̣̣̣̣̣	𑀓̣̣̣̣̣	𑀓̣̣̣̣̣	𑀓̣̣̣̣̣	𑀓̣̣̣̣̣	𑀓̣̣̣̣̣	𑀓̣̣̣̣̣	𑀓̣̣̣̣̣					

TABLE 3.2 – A
Usually pronounced /ɔ/ or /a/

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
ạ (back a)	ᨾ	ᩉ᩠ᨦ	ᩉ᩠ᨦ	กอ	ກອ	ᨾᩉ	ᩉ᩠ᨦ	ᩉ᩠ᨦ	ᩉ᩠ᨦ	ᩉ᩠ᨦ	ᩉ᩠ᨦ	ᩉ᩠ᨦ	ᩉ᩠ᨦ		
aa (a+ạ)			ᩉ᩠ᨦ			ᩉ᩠ᨦ	ᩉ᩠ᨦ								
ắ (ạ + ́) (<i>musikadānt</i> ; inherent vowel)			ᩉ᩠ᨦ	ก๋อ											
ắạ (ạ + ́ + ́)			ᩉ᩠ᨦ												
ắv (ắ+v)				กัว											
ằ (ạ + ̀) (Lao/Lanna diacritical ằ; Mon ằ as glottal stop or final h)					ᨾ	ᨾ	ᨾ	ᨾ						ᨾ	
ằ (ạ+ằ) (Tham ằ with diacritical ằ)						ᨾ	ᨾ	ᨾ	ᨾ	ᨾ					
ằ (modified ằ) (Khmer/Khom <i>daṇḍaghāt</i> as ằ; Khuen/Lue diacritical ằ)	ᨾ	ᨾ	ᨾ					ᨾ	ᨾ	ᨾ					
ằ (ạ+ằ)								ᨾ	ᨾ	ᨾ					
ằạ (ạ+ằ)											ᨾ				
ằm												ᨾ			

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
ay (a+y, ̣ vowel)	ក័យ	កយ													
ay (a+y, ̣ vowel)	ក័យ	កយ													
ay (a+y, ̣ vowel)						໋	໋	໋	໋	໋	໋				
ay (̣+y)			ກຸຍ	ກອຍ	ກອຍ	ກອຍ									
ay (̣+y)			ກຸຍ	ກອຍ	ກອຍ	ກອຍ	ກຸຍ	ກຸຍ	ກຸຍ	ກຸຍ					
ay (̣+y)												ກຸຍ	ກຸຍ		
aay (ạ+y)			ກຸຍ												
̣	ក	ក	ក												
̣ (̣ written like ū)			ກຸຍ												
̣ (̣+̣)			ກຸຍ												
̣ (̣+̣)					ກອ		ກຸຍ	ກຸຍ	ກຸຍ	ກຸຍ					
̣ (̣+̣)	ກ	ກ	ກ	ກ											
̣ (̣+̣)			ກ	ກ				ກ							

TABLE 3.3 — Ā

Usually pronounced /a:/, sometimes /ia/ in Khmer or /ɛ̃a/ in Mon

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
ā (long a)	កា	ကာ	กา	กา	ກາ	ᨠᩅ	ᩅ᩠ᨦ	ᩅ᩠ᨦ	ᩅ᩠ᨦ	ᩅ᩠ᨦ	ᩅ᩠ᨦ	ᩅ᩠ᨦ	ᩅ᩠ᨦ	ကာ	ကာ
ām (ā+m)	កាំ	ကံ	กัม	กัม	ກံ	ᨠ᩵	᩠᩵ᨦ	᩠᩵ᨦ	᩠᩵ᨦ	᩠᩵ᨦ				ကံ	
āḥ (ā+h)	កាះ	ကား													
āy (ā+y)	កាយ	ကယ	กาย	กาย	ກາຍ	ᨠ᩶									
āy (ā+y)	ကျ	ကျ	ကျ		ກຽ	ᨠ᩶					ᩅ᩠ᨦ				
āỹ (ā+ỹ)													ခံ	ကဲ	
āv (ā+v)	ကာវ	ကာ	กา	กา	ກາວ	ᨠ᩶									
āṿ (ā+ṿ)													ကဝ်		
āỵ (ā+ỵ)	ကျ	ကျ	ကျ												
ǎ (ā+ǎ) (final āk)	ကံ	ကံ					ᩅ᩠ᨦ	ᩅ᩠ᨦ	ᩅ᩠ᨦ	ᩅ᩠ᨦ					

ONLINE APPENDIX

TABLE 3.4 — I

Usually pronounced /i/, various other pronunciations such as /e/, /ə/, /ɛi/, or /ɜ/

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
i	ᨾ	ᨾ᩵	ᨾ᩵	ᨾ᩵	ᨾ᩵	ᨾ᩵	ᨾ᩵	ᨾ᩵	ᨾ᩵	ᨾ᩵		ᨾ᩵	ᨾ᩵	ᨾ᩵	ᨾ᩵
iḥ (i+h)	ᨾ᩠᩵	ᨾ᩠᩵													
iṁ (i+ ṁ)	ᨾ᩠᩵᩵	ᨾ᩠᩵᩵	ᨾ᩠᩵᩵	ᨾ᩠᩵᩵	ᨾ᩠᩵᩵		ᨾ᩠᩵᩵	ᨾ᩠᩵᩵	ᨾ᩠᩵᩵	ᨾ᩠᩵᩵		ᨾ᩠᩵᩵		ᨾ᩠᩵᩵	ᨾ᩠᩵᩵
iṁ (i+ modified ṁ)														ᨾ᩠᩵᩵	
iv (i+v)			ᨾ᩠᩵᩵	ᨾ᩠᩵᩵	ᨾ᩠᩵᩵	ᨾ᩠᩵᩵									
iṿ												ᨾ᩠᩵᩵			
iṽ (i+ṿ)			ᨾ᩠᩵᩵												
ī (long i)	ᨾ᩠᩵᩵	ᨾ᩠᩠᩵᩵᩵	ᨾ᩠᩠᩵᩵᩵	ᨾ᩠᩠᩵᩵᩵	ᨾ᩠᩠᩵᩵᩵	ᨾ᩠᩠᩵᩵᩵	ᨾ᩠᩠᩵᩵᩵	ᨾ᩠᩠᩵᩵᩵	ᨾ᩠᩠᩵᩵᩵	ᨾ᩠᩠᩵᩵᩵	ᨾ᩠᩠᩵᩵᩵	ᨾ᩠᩠᩵᩵᩵	ᨾ᩠᩠᩵᩵᩵	ᨾ᩠᩠᩵᩵᩵	ᨾ᩠᩠᩵᩵᩵
īḥ (ī+h)											ᨾ᩠᩠᩵᩵᩵				

TABLE 3.5 – I

Usually pronounced /i/ or /ɯ/

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon
i	ᨾ᩵	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ				
iḥ (i+h)	ᨾ᩠᩵ᨿ	ᨾᩴ᩠ᨿ									ᨾᩴ᩠ᨿ			
ī (long i)	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ			
īḥ (ī+h)	ᨾᩴ᩠ᨿ	ᨾᩴ᩠ᨿ						ᨾᩴ᩠ᨿ	ᨾᩴ᩠ᨿ	ᨾᩴ᩠ᨿ				
īy (ī+y)											ᨾᩴ᩠ᨿ			
īa (ī+a)				ᨾᩴ										

TABLE 3.6 – U

Usually pronounced /u/, sometimes also /ɔ/, /o/ or /ɔo/

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
u	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ	ᨾᩴ
uṁ (u+m)	ᨾᩴ᩠ᨿ	ᨾᩴ᩠ᨿ		ᨾᩴ᩠ᨿ	ᨾᩴ᩠ᨿ							ᨾᩴ᩠ᨿ		ᨾᩴ᩠ᨿ	ᨾᩴ᩠ᨿ
uḥ (u+h)	ᨾᩴ᩠ᨿ	ᨾᩴ᩠ᨿ													
uy (u+y)	ᨾᩴ᩠ᨿ	ᨾᩴ᩠ᨿ		ᨾᩴ᩠ᨿ	ᨾᩴ᩠ᨿ	ᨾᩴ᩠ᨿ									
uṽ (u+ỹ)														ᨾᩴ᩠ᨿ	
uy· (u+y·)												ᨾᩴ᩠ᨿ			

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
uv (u+v)	ꠘꠞ	ꠘꠞ													
ũ (u+ă, final uk)	ꠘꠞ	ꠘꠞ					ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ					
ū (long u)	ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ
ūḥ (ū+h)	ꠘꠞ	ꠘꠞ													
ũ̄ (ū+ă, final ūk)	ꠘꠞ	ꠘꠞ					ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ					
ūy (ū+y)											ꠘꠞ				
ūv (ū+v)	ꠘꠞ	ꠘꠞ													
û (ū+â)							ꠘꠞ	ꠘꠞ	ꠘꠞ	ꠘꠞ					

TABLE 3.7 — R̥

Vocalic consonants, borrowed from Sanskrit

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
r̥	ꠘꠞ	ꠘꠞ	ꠘꠞ												ꠘꠞ
r̄ (long r̥)	ꠘꠞ	ꠘꠞ	ꠘꠞ												ꠘꠞ
l̥	ꠘꠞ	ꠘꠞ													ꠘꠞ
l̄ (long l̥)	ꠘꠞ	ꠘꠞ													ꠘꠞ

TABLE 3.8 — E
Usually pronounced /e/, sometimes /ɛe/ or /ɛ/

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
e	ເກ	ເ	เ	เ	เ	เ	เ	เ	เ	เ	เ	เ	เ	เ	เ
ě (e+ǎ)														ကဲ	
em̐ (e+m̐)														ကဲ	
eh̐ (e+h̐)	ເ	ເ	เ	เ	เ	เ	เ	เ	เ	เ	เ	เ			
eyy	ເຍ	ເ	เ				เ	เ	เ	เ	เ				
eÿ (e+ÿ)														ကဲ	
ev (e+v)				เ	เ	เ									
eṽ (e+ṽ)												เ			
ě (modified e) (Shan diacritical e)													เ		
ě (e+ǎ)					เ	เ									
ěv (ě+v)					เ	เ									
ě (e+ǎ)				เ				เ							
ěv (ě+v)				เ											

ONLINE APPENDIX

TABLE 3.9 — Æ
Usually pronounced /ɛ/, sometimes /æ/

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
æ	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵		ᨠ᩵ᩁᩬ᩵		
æḥ (æ+h)			ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵					ᨠ᩵ᩁᩬ᩵				
æv (æ+v)	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵			ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵									
æṽ (æ+v̄)	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵				ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵					
ǣ (æ+ǣ)					ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵									
ǣ (modified æ) (Tham/Shan diacritical æ)							ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵			ᨠ᩵ᩁᩬ᩵		
ae (variant æ)	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵						ᨠ᩵ᩁᩬ᩵							
aeṽ (ae+v̄)	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵													
aeṽ (ae+v̄)	ᨠ᩵ᩁᩬ᩵	ᨠ᩵ᩁᩬ᩵													

TABLE 3.10 — AI

Usually pronounced /ai/, sometimes as /au/, /ɜi/, or /aw/

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
ai	ไ	ไ	ไ	ไ	ไ	ไ	ไ	ไ	ไ	ไ	ไ				ကဲ
aiḥ (ai +h)	ไḥ	ไḥ													
aiy (ai+y)	ไย	ไย		ไย											
aiỵ (ai+ỵ)	ไย̣	ไย̣	ไย̣				ไย̣	ไย̣	ไย̣	ไย̣					
ai̋ (modified ai, conceived as ai+i)				ไ	ไ	ไ	ไ	ไ	ไ						
aỵ (a+ỵ)												ကယ်		ကယ်	ကယ်
ay̋ (a+y̋)												က'	က'	က'	
ǎy (ǎ+y)	ไ	ไ	ไ	ไ	ไ	ไ	ไ	ไ	ไ	ไ					
ǎỵ (ǎ+ỵ)	ไ	ไ	ไ		ไ	ไ									

TABLE 3.11 – O
Usually pronounced /o/, sometimes as /aw/, /ao/ or /ɔ/

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
o	ເກ	ເກ	เก	เกา			ເວ	ເວ	ເວ	ເວ		၇/၈	ဂေ	ကော	ကော
om̐ (o+m̐)														ကောံ	
oh̐ (o+h̐)	ເກး	ເကး	เกะ	เกาะ	เกาะ	เกาะ									
ov (o+v)	ເກវ	ເກវ													
oy (o+y)	ເກຍ	ເກຍ													
oỵ (o+ỵ)	ເကျ	ເကျ	เกฺย												
oy̥ (o+y̥)														ကော့	
öy̥ (ö+y̥)			เกฺย												
ô (o+ â) (Lanna mái kô kò)					ເກົາ	ເກົາ	ເດົາ	ເດົາ	ເດົາ	ເດົາ					
ó (modified o) (Lanna mái ko hâ hnín)							ເດ	ເດ							
õ							ເດ	ເດ	ເດ	ເດ					
ō (Tai o)			เกอ	โก	โก	โ	โ	โ	โ	โ	โ				
ōỵ (ō+ỵ)					โ	โ									

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
ōy (ō+y)					ໂຍ	ຍ້					လၢ				
ô (ō+â)							ໂ	ໂ	ໂ	ໂ					
ōh (ō+h)			ໂ	ໂ	ໂ	ໂ					ໂ				
ō (ō+a)							ໂ	ໂ	ໂ	ໂ					
ōh (ō+h)							ໂ	ໂ	ໂ	ໂ					

TABLE 3.12 — AU

Usually pronounced /aw/, /au/, or /ai/, sometimes as /əi/

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
au	က	က	See O	See O	See O	See O	က	က	က	က			က		က
auh (au+h)	က	က													
auv (au+v)	က	က													
av (a+v)												က			
äv (a+v̄)												က		က	
äv̄m (äv̄+m)												က			
ay (a+y)													က		
äv̄ (a+v̄)												က			

TABLE 3.13 — Ǽ

Usually pronounced /ɤ/ or /ə/, sometimes /aə/ in Khmer

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
ǽ	ເກີ	ເកີ	เกิ	เกิ	ເກີ	ເກີ	ເກີ	ເກີ	ເກີ	ເກີ					
ǽy (ǽ+y)	ເກີຍ	ເກີຍ	เกีย												
ǽy (ǽ+y)	ເກີງ	ເກີງ	เกีย												
ǽa (ǽ+a)			เกีย												
ǽa (ǽ+a)			เกีย	เกิ			ເກີ	ເກີ	ເກີ	ເກີ					
ǽah (ǽ+a+h)								ເກີ	ເກີ	ເກີ					
ǽay (ǽ+a+y)			เกีย				ເກີ	ເກີ	ເກີ	ເກີ					
ǽay (ǽ+a+y)			เกีย				ເກີ	ເກີ	ເກີ	ເກີ					
ǽ (ǽ+i)			เกิ												
ǽa (ǽ-i+a)			เกิ	เกิ				ເກີ							
ǽa (ǽ+a+a)			เกิ												
ǽah (ǽ+a+h)				เกิ				ເກີ							
ǽy (ǽ+y)				เกิ											

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
ø _y (ø+y)			เอย												
ø _y (ø+ǎ+y)			เอย												
ø _a (ø+''+a)			เอย	เอย											
ǣ (ǣ+ī)	เอย	เอย	เอย		เอย	เอย					เอย				
ǣ _h (ǣ+h)	เอย	เอย									เอย				
ǣ _y (ǣ+y)	เอย	เอย			เอย	เอย									
ǣ _y (ǣ+y)	เอย	เอย			เอย	เอย					เอย				
ǣ (ǣ+ī)	เอย	เอย							เอย	เอย					
ǣ _y (ǣ+y)	เอย	เอย													
ǣ _y (ǣ+y)	เอย	เอย							เอย	เอย					

TABLE 3.14 — ĆA

Usually pronounced /wa/ or /iə/; monophthongized in Khuen and Lue as /ʌ/ or /ə/

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
œa (œ+a)			เฝ้า												
œa (œ+a)				เฝ้า			เฝ้า	เฝ้า	เฝ้า	เฝ้า					
œaa (œ+a+a)			เฝ้า				เฝ้า	เฝ้า	เฝ้า	เฝ้า					
œaah (œ+a+a+h)							เฝ้า	เฝ้า	เฝ้า	เฝ้า					
œa (œ+a)			เฝ้า												
œa (œ+a)			เฝ้า	เฝ้า											
œa (œ+a)					เฝ้า	เฝ้า	เฝ้า	เฝ้า	เฝ้า	เฝ้า					
œaah (œ+a+a+h)									เฝ้า						
œay (œ+a+y)							เฝ้า	เฝ้า	เฝ้า	เฝ้า					
œay (œ+a+y)							เฝ้า	เฝ้า	เฝ้า	เฝ้า					
œa (œ+a)			เฝ้า												
œa (œ+a)	เฝ้า	เฝ้า	เฝ้า	เฝ้า	เฝ้า	เฝ้า	เฝ้า	เฝ้า	เฝ้า	เฝ้า					
œaa (œ+a+a)							เฝ้า	เฝ้า	เฝ้า	เฝ้า					

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
ṭṭaa• (ṭṭ+a+a+•)							ເໝີຂໍ	ເໝີວຸ້	ເໝີວຸ້						
ṭṭaaḥ (ṭṭ+a+a+ḥ)								ເໝີຂຸ	ເໝີວຸຸ	ເໝີວຸຸ					
ṭṭaaḥ• (ṭṭ+a+a+ḥ+•)								ເໝີຂຸ	ເໝີວຸຸ	ເໝີວຸຸ					
ṭṭaḥ (ṭṭ+a+ḥ)			ເໝີ	ເໝີ											
ṭṭay (ṭṭ+a+y)	ເໝີ	ເໝີ	ເໝີ	ເໝີ	ເໝີ	ເໝີ									
ṭṭay (ṭṭ+a+y)					ເໝີ	ເໝີ	ເໝີ	ເໝີ	ເໝີ	ເໝີ					
ṭṭv (ṭṭ+v)				ເໝີ											
ṭṭv (ṭṭ+v)				ເໝີ											
ṭṭv (ṭṭ+ " + v)				ເໝີ											
ṭṭv (ṭṭ+ " + v)				ເໝີ											

TABLE 3.15 — IA

Usually pronounced ie/ or /iə/; if ya is serving as a consonant (/j/), transliterate as ya

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
ia	ꨀꩃ	ꨀꩃ	ꨀꩃ		ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ					
ia ^v							ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ					
ia ^{v•}								ꨀꩃ	ꨀꩃ	ꨀꩃ					
ĩa (ia+ [˘])					ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ					
ĩa	ꨀꩃ	ꨀꩃ													
ie (ia+e)	ꨀꩃ	ꨀꩃ			ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ					
ie _h (ie+h)	ꨀꩃ	ꨀꩃ					ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ					
iey (ie+y)	ꨀꩃ	ꨀꩃ													
iev (ie+v)	ꨀꩃ	ꨀꩃ													
ĩe (i+ [˘])					ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ					
ĩe _h (ie+ [˘] +h)							ꨀꩃ	ꨀꩃ	ꨀꩃ	ꨀꩃ					
īe (ie+i)			ꨀꩃ												
īe (ie+ī)			ꨀꩃ	ꨀꩃ											
īe _h (īe+h)				ꨀꩃ											

TABLE 3.16 — UA

Usually pronounced /ua/ or /uə/; if *va* is serving as a consonant (/v/ or /w/), transliterate as *va*

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
ua	ກູ້	ກູ	กู	กา	ກວ	ວາ									
ũa (ua+ă)			กั๊	ก้า	ກັວ	ວັວ									
ũaḥ (ua+ă+ḥ)			กั๊ះ	ก้าะ											
ũa (ua+â)					ກົວ	ວົວ		ຜູ້	ກູ້	ກູ້					
ũaḥ (ua+â+ḥ)					ກົວະ	ວົວະ		ຜູ້ະ	ກູ້ະ	ກູ້ະ					
uay (va+y)			กวย	กาย	ກວຍ	ວາຍ									
uay (va+y)			กฺ				ຜູ້	ຜູ້	ກູ້	ກູ້					
uo (modified ua)	ກູ້	ກູ													
uö (uo+ă)	ກູ້:	ກູ:													
uoḥ (uo+ḥ)	ກູ້:	ກູ:													
uoy (uo+y)	ກູ້ຍ	ກູ້ຍ													
uoy (uo+y)	ກູ້ຢ	ກູ້ຢ													

TABLE 3.17 — UI

Various pronunciations, including /o/, /i/, /v/, and /ɜ̃/

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
ui (u+i) (O. Khmer <i>i/ī</i> ; Lanna <i>mái ko cú cǐ</i>)	ᨾ᩵	᩠᩵ᩅ						᩠᩵ᩅ				᩠᩵ᩅ	᩠᩵ᩅ	᩠᩵ᩅ	᩠᩵ᩅ
uim̄ (u+iṁ)														᩠᩵ᩅ	
uiỵ (ui+ỵ)												᩠᩵ᩅ	᩠᩵ᩅ		
uiṽ (ui+ṽ)														᩠᩵ᩅ	
uiṿ (ui+ṿ)												᩠᩵ᩅ	᩠᩵ᩅ		
uiṽ (ui+ṽ)														᩠᩵ᩅ	
ūi (ū+i)													᩠᩵ᩅ		
ūī (ū+ī)									᩠᩵ᩅ						

TABLE 4 — Tone Marks

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Burm.
No tone marker (e.g. <i>kā</i>) (also Modern Shan tone 1)	ᦅᦺ	ᦅᦺ	ᦅᦺ	กา	ກາ	ᦅᦺ	ᦅᦺ	ᦅᦺ	ᦅᦺ	ᦅᦺ	ᦅᦺ	ᦅᦺ	ᦅᦺ	ᦅᦺ	ᦅᦺ
^ˊ (e.g. <i>kā̊</i>) (<i>mái ek</i> ; Modern Shan tone 2)	ᦅᦺᦺ	ᦅᦺᦺ	ᦅᦺᦺ	กา̊	ກາ̊	ᦅᦺᦺ	ᦅᦺᦺ	ᦅᦺᦺ	ᦅᦺᦺ	ᦅᦺᦺ	ᦅᦺᦺ		ᦅᦺᦺ		
^ˋ (e.g. <i>kǎ̋</i>) (<i>mái dō</i> ; Modern Shan tone 3)	ᦅᦺᦺ	ᦅᦺᦺ	ᦅᦺᦺ	กา̋	ກາ̋	ᦅᦺᦺ	ᦅᦺᦺ	ᦅᦺᦺ	ᦅᦺᦺ	ᦅᦺᦺ	ᦅᦺᦺ		ᦅᦺᦺ		
^{ˊˊ} (e.g. <i>kǎ̊̊</i>) (doubled <i>mái dō</i> , with or without vowel change)							ᦅᦺᦺ	ᦅᦺᦺ	ᦅᦺᦺ	ᦅᦺᦺ					
^{ˊˊ} (e.g. <i>kǎ̊̊</i>) (<i>mái trī</i> ; Khuen <i>mái kô hnṵṵ</i> ; Modern Shan tone 4)			ᦅᦺᦺ	กา̊̊	ກາ̊̊				ᦅᦺᦺ	ᦅᦺᦺ			ᦅᦺᦺ		
^{ˋˋ} (e.g. <i>kǎ̋̋</i>) (<i>mái cǎivā</i> ; Khuen <i>mái sañ hnṵṵ</i>)			ᦅᦺᦺ	กา̋̋	ກາ̋̋				ᦅᦺᦺ	ᦅᦺᦺ					
^{ˊˊ} (e.g. <i>kā̊̊</i>) (Khuen <i>mái sām hnṵṵ</i> ; Modern Shan tone 5)									ᦅᦺᦺ	ᦅᦺᦺ			ᦅᦺᦺ		
^{ˊˊ} (e.g. <i>kā̊̊</i>) (Modern Shan tone 6)													ᦅᦺᦺ		
ᦅ (e.g. <i>kāḥ</i>) (Old Shan tone 4/Burmese high tone marker)													ᦅᦺᦺ		ᦅᦺᦺ
^ᦅ (e.g. ... <i>yaᦅ</i>) (<i>a</i> used as glottal stop/tone marker)														ᦅᦺᦺ	ᦅᦺᦺ
^ᦅ (e.g. ... <i>yᦅ</i>) (subscript <i>a</i> used as glottal stop/tone marker)															ᦅᦺᦺ
˙ (e.g. <i>caṁ.</i> / <i>kaṁ.</i>) (Old Shan tone 5/Burmese creaky tone mark)													ᦅᦺᦺ		ᦅᦺᦺ

TABLE 5 — Other Diacritics

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	La o	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
_ (macron below) (For marking unexpected subscripts)															
◌ (ring below) (Thai <i>bindu</i> , for marking consonant prior to a subscript)				๑											
˘ (macron above) (Thai <i>yāmakār</i> , older alternative to Thai <i>bindu</i>)				๑̄											
· (interpunct) (<i>virāma</i> for Old Khmer, Lue, Shan, Mon, Burmese)	ក្រី	နီ									လွ	၎်	၎်		၎်
• (bullet) (Khmer/Khom <i>daṇḍaghāt</i> , Tham <i>ra hām</i>)	ក្រី	နီ	နီ	ก	ṅ	၎်		၎်	ကီ	ကီ					
` (grave above) (Khmer <i>rassasaññā</i> , placed over final consonant; Thai <i>phan dān</i> , syllable marker)	ក̀	န̀	န̀	ก̀											
˙ (hook above) (Khmer <i>kākapād</i> , placed over consonant for emphasis/sound change)	ക̣	န̣													
' (apostrophe) (Khmer <i>avagraha</i> , Sanskrit only)	ṣ	ṣ													
~ (tilde) (Tham <i>mái jāṃ</i> , syllable/word repeater mark; Lik Tai sometimes uses a double <i>virāma</i> ~ instead)								၎်	က၎်	က၎်			၎်		
& (Khmer <i>lekh do</i> , Thai <i>mái yamak</i> ; word/phrase repeater mark)	๗	၎	၎	๗	၎	၎		၎							

TABLE 6 — Numerals

Alternative sets may be differentiated by underlining; digits rarely attested for Lik Tai/Tai Ahom numbers 3 [sam], 4 [sī], and 5 [hā]

Trans.	Atta:	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna 1	Khuen 1	T. Lue 1	Lanna 2	Khuen 2	T. Lue 2	Lue	Lik	Shan	Mon	Brm.
0	๐	၀	၀	๐	๐	໐	໐	໐	໐	໐	໐	໐	໐	໐	໐	໐	၀	၀	໐
1	๑	၁	၁	๑	๑	໑	໑	໑	၎်	၎်	၎်	၎်	၎်	၎်	၁/င	၁	၁	၁	၁
2	๒	၂	၂	၂	၂	໒	໒	໒	၎်	၎်	၎်	၎်	၎်	၎်	၂	၂	၂	၂	၂
3	๓	၃	၃	ၓ	ၓ	໓	໓	໓	၎်	၎်	၎်	၎်	၎်	၎်	၃	၃	၃	၃	၃
4	๔	၄	၄	ၔ	ၔ	໔	໔	໔	၎်	၎်	၎်	၎်	၎်	၎်	၄	၄	၄	၄	၄
5	๕	၅	၅	ၕ	ၕ	໕	໕	໕	၎်	၎်	၎်	၎်	၎်	၎်	၅	၅	၅	၅	၅
6	๖	၆	၆	ၖ	ၖ	໖	໖	໖	၎်	၎်	၎်	၎်	၎်	၎်	၆	၆	၆	၆	၆
7	๗	၇	၇	ၗ	ၗ	໗	໗	໗	၎်	၎်	၎်	၎်	၎်	၎်	၇	၇	၇	၇	၇
8	๘	၈	၈	ၘ	ၘ	໘	໘	໘	၎်	၎်	၎်	၎်	၎်	၎်	ၘ	ၘ	ၘ	ၘ	ၘ
9	๙	၉	၉	ၙ	ၙ	໙	໙	໙	၎်	၎်	၎်	၎်	၎်	၎်	ၙ	ၙ	ၙ	ၙ	ၙ

TABLE 7 — Abbreviations

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
@ (Khom/Thai <i>ăn vằ</i> or <i>ăn vằ</i> , cf. Khmer <i>rī</i>)			๙				ຯ	ຯ	ຯ	ຯ					
lä (Tham <i>læ</i> abbreviation)							ລ	ລ	ລ	ລ	ລ				
læ (la+æ) (Tham <i>læ</i> with diacritical <i>æ</i>)							ລີ	ລີ	ລີ	ລີ					
læ (lä+æ) (Tham <i>læ</i> abbreviation with normal <i>æ</i>)							ເລ	ເລ	ເລ	ເລ					
læ (lä+æ) (Tham <i>læ</i> with diacritical <i>æ</i> and normal <i>æ</i>)							ເລີ	ເລີ	ເລີ	ເລີ					
läv (Tham <i>læ</i> abbreviation with final <i>v</i>)							ໂວ	ໂວ	ໂວ	ໂວ	ໂວ				
läv (Tham <i>l</i> with diacritical <i>æ</i> and final <i>v</i>)							ໂວ	ໂວ	ໂວ	ໂວ					
3dī (Khmer/Khom “3 times” marker, Thai <i>sām dī</i>)	๓	๓	๓												
3dhī (Khmer/Khom “3 times” marker, Thai <i>sām dī</i>)	๓	๓	๓												
` (backtick) (Generic special character mark)															
`e															၏
`l															၎
`n															ဏ်

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
`r													၍		၍
`! (Shan/Lik Tai vocative/exclamation mark)												၍	၍		
`jǎñ								ဝ							
`viañ								⊕							
`viañ-hvāk								⊕							
`svār								သ							

TABLE 8 — Punctuation

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
 (Single <i>daṇḍa</i>)															
 (Double <i>daṇḍa</i>)															
 (Khmer <i>khaṇḍ</i> , Thai <i>ǎñ-gǎn tiēv</i>)	၍	၍	၍	၍	၍	၍	၍	၍	၍	၍					
# (Variant of Khmer <i>khaṇḍ</i> , Thai <i>ǎñ-gǎn tiēv</i>)	၍	၍	၍	၍	၍	၍	၍	၍	၍	၍					
 (Khmer <i>rapaḥ caḥ</i> , Thai <i>ǎñ-gǎn tiēv</i>)	၍	၍	၍	၍			၍	၍	၍	၍					
# (Variant of Khmer <i>rapaḥ caḥ</i> , Thai <i>ǎñ-gǎn tiēv</i>)	၍	၍	၍	၍			၍	၍	၍	၍					

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
la (Khmer <i>peyyāl</i> , Thai <i>paiyāl hñàì</i>)	១ល ១	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ					
pa (Variant of Khmer <i>peyyāl</i> , Thai <i>paiyāl hñàì</i>)	១ប ១	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ				ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ					
pe (Variant of Khmer <i>peyyāl</i> , Thai <i>paiyāl hñàì</i>)	១បេ ១	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ				ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ					
◎ (Khmer <i>bhnaek mǎn</i> , Thai <i>bạn mǎn</i>)	◎	◎	◎	◎			◎	◎	◎	◎					
◎ (Thai <i>bạn mǎn bǎn hnū</i>)	◎	◎	◎	◎								◎			
○ (Thai <i>bạn tǎn</i>)				○											
% (Khmer <i>go mǔtr</i> , Thai <i>gō mǔtr</i>)	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ			ᦺᦺᦺᦺ	ᦺᦺᦺᦺ							
% (Variant of <i>go mǔtr</i>)	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ												
:% (Variant of <i>go mǔtr</i>)	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ	ᦺᦺᦺᦺ											
÷ (Khmer colon, <i>camñuc bīr gūs</i>)	÷	÷	÷												
§ (Tham punctuation mark)								§							
⌘ (Section mark)	⌘	⌘	⌘					⌘							
* (Floral section mark)								*							

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
_ (underscore) (Space in original text)															
/ (Line break in original text)															
○ (String hole break in original text)															
# (Space filler prior to line break)	:	:	๕				◦:	◦:	◦:	◦:					
≤ (less-than or equal sign) (Bracketed list, repeated for each item)	{	{	{	{	{		{								
≥ (greater than or equal sign) (End of bracketed list, generally unmarked)															

TABLE 9 — Editorial Marks

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
(unmarked in original) (Editorial space to mark word boundaries)															
- (hyphen) (Hiatus bridge for line breaks, hole breaks, and vowel hiatus)															
= (equal sign) (For separating pairs of words joined with a subscript)															
° (degree symbol) (Generic abbreviation marker)															
≈ (double tilde) (Tham compounds without <i>mái jām</i>)															

ONLINE APPENDIX

Roman	Kh. Chr.	Kh. M.	K. Thai	Thai	Lao	Lao B.	T. Lao	Lanna	Khuen	T. Lue	Lue	Lik	Shan	Mon	Brm.
< >															
(Left and right edges of syllable divided by a break)															
* (asterisk)															
(Non-attested/hypothetical form marker)															
× (multiplication sign)															
(Illegible letter/ <i>aksara</i>)															
[[]]															
(Scribal restoration of illegible portion)															
{ { } }															
(Scribal deletion of erroneous portion)															
((↑))															
(Scribal insertion above)															
((↓))															
(Scribal insertion below)															
^ (circumflex accent)	' ក	' ម	' ឃ												
(Annotator's inserted space above)															
ˆ (caret)	ក	ម	ឃ												
(Annotator's word isolation marker below)															
()															
(Generic editorial addition)															
[]															
(Editorial restoration of illegible portion)															
{ }															
(Editorial deletion of erroneous/redundant portion)															